

# Boolean Irreducibility and Phase Transitions in $NK$ -Kauffman Networks

Federico Zertuche

Instituto de Matemáticas, Unidad Cuernavaca  
Universidad Nacional Autónoma de México  
A.P. 273-3, 62251 Cuernavaca, Mor., México.  
zertuche@matcuer.unam.mx

## Abstract

In [B. Derrida, and D. Stauffer, *Phase Transitions in Two-Dimensional Cellular Automata*. Europhys. Lett. **2** (1986) 739], a mean field study of the  $NK$ -Kauffman Networks was done. They obtained a phase transition curve  $K_c 2p(1-p) = 1$  ( $0 < p < 1$ ) for the critical average connectivity  $K_c$  in terms of the bias  $p$  of extracting a “1” for the output of the automata. Values of  $K$  bigger than  $K_c$  correspond to the so-called chaotic phase; while  $K < K_c$ , to an ordered phase. In [F. Zertuche, *On the robustness of  $NK$ -Kauffman networks against changes in their connections and Boolean functions*. J. Math. Phys. **50** (2009) 043513], a new classification for the Boolean functions, called *Boolean irreducibility* permitted the study of new phenomena of  $NK$ -Kauffman Networks. In the present work I study, once again the mean field treatment for  $NK$ -Kauffman Networks, correcting it for *Boolean irreducibility*. A new phase transition curve is founded. In particular, for  $p = 1/2$  the predicted value  $K_c = 2$  by Derrida & Stauffer changes to  $K_c = 2.62140224613\dots$

**Short title:** *Boolean Irreducibility and Phase Transitions*

**Keywords:** *Cellular automata, Boolean irreducibility, binary functions, phase transitions,  $NK$ -Kauffman Networks.*

**PACS numbers:** 87.10.-e, 87.10.Mn, 87.10.Ca, 05.70.Fh

## 1. Introduction

$NK$ -Kauffman networks have been widely studied due to its applications in theoretical biology; specially they are useful for the study of the genotype-phenotype map  $\Psi$  [1–4]. As long as they are randomly constructed, they are well suited for study with statistical techniques; particularly by the use of the mean field approximation.

A  $NK$ -Kauffman network consists of  $N$  Boolean variables  $S_i(t) \in \mathbb{Z}_2$  ( $i = 1, \dots, N$ ), which evolve deterministically in discrete time  $t = 0, 1, 2, \dots$  according to Boolean functions on  $K$  ( $0 \leq K \leq N$ ) of these variables at the previous time  $t - 1$ . For every site  $i$ , a  $K$ -Boolean function  $\mathfrak{F}_i : \mathbb{Z}_2^K \rightarrow \mathbb{Z}_2$  is chosen randomly and independently with a bias probability

$$0 < p < 1, \quad (1)$$

that  $\mathfrak{F}_i = 1$ ; for each of its possible  $2^K$  arguments. Also, for every site  $i$ ,  $K$  inputs (the connections) are randomly selected from a uniform distribution among the  $N$  Boolean variables of the network, without repetition. So, for each site  $i$ , and each extraction  $E$  ( $0 \leq E \leq K - 1$ ); an input  $j$  ( $1 \leq j \leq N$ ) is extracted with probability

$$P_E(j) = \begin{cases} \frac{1}{N-E}, & \text{if } j \text{ has not been previously extracted} \\ 0, & \text{if } j \text{ has already been extracted} \end{cases} \quad (2)$$

Once the  $K$  inputs and the functions  $\mathfrak{F}_i$  have been selected, a Boolean deterministic  $NK$ -Kauffman network has been defined. It evolves deterministically, and synchronously in time, according to the rules

$$S_i(t+1) = \mathfrak{F}_i(S_{i_1}(t), S_{i_2}(t), \dots, S_{i_K}(t)), \quad i = 1, \dots, N, \quad (3)$$

where  $i_m \neq i_n$ , for all  $m, n = 1, 2, \dots, K$ , with  $m \neq n$  since, from (2), each input is different.  $NK$ -Kauffman networks are then a special type of all the Boolean endomorphisms  $\mathcal{B}_N$ ;

$$\mathcal{B}_N = \{f : \mathbb{Z}_2^N \longrightarrow \mathbb{Z}_2^N\}.$$

Let  $\mathcal{L}_K^N$  be the set of  $NK$ -Kauffman networks. Then  $\mathcal{L}_K^N \subseteq \mathcal{B}_N$ , and  $\mathcal{L}_N^N \cong \mathcal{B}_N$  [3,4]. In Ref. [3], a study of the injective properties of the map

$$\Psi : \mathcal{L}_K^N \rightarrow \mathcal{B}_N \quad (4)$$

was pursuit. Using the fact that  $\mathcal{B}_N \cong \mathcal{G}_{2^N}$ , where  $\mathcal{G}_{2^N}$  is the set of functional graphs on  $2^N$  points (*i.e.* the directed graphs with **out-degree one**, and loops allowed) [5]; the average number  $\vartheta(N, K)$  of elements in  $\mathcal{L}_K^N$  that  $\Psi$  maps into the same Boolean function was calculated [3,4]. The results showed that there exists a critical average connectivity  $\hat{K}$  in the region  $K \sim \mathcal{O}(\ln \ln N)$  for  $N \gg 1$ , given by

$$\hat{K} \approx \log_2 \log_2 \left( \frac{2N}{\ln 2} \right) + \mathcal{O} \left( \frac{1}{N \ln N} \right); \quad (5)$$

such that  $\vartheta(N, K) \approx e^{\varphi N} \gg 1$  ( $\varphi > 0$ ) or  $\vartheta(N, K) \approx 1$ , depending on whether  $K < \hat{K}$  or  $K > \hat{K}$ , respectively.

An important challenge, since the proposal by Kauffman of their networks [2], has been the analytic calculation of their average dynamics in terms of the network parameters:  $N, K$ , and  $p$ ; their number, their connectivity, and their extraction bias, respectively. Until now only some special cases have been completely solved: The case  $K = N$ , the so-called *random map model* [5,6], and the extreme case  $K = 1$  [7].

In 1986, Derrida & Stauffer [8] tried a mean field approach (where  $N \rightarrow \infty$ ) to study the problem. They showed that there exists a critical connectivity  $K_c$ , such that the Hamming distance between two nearby states grows or decays exponentially according to whether  $K > K_c$  or  $K < K_c$ , respectively; where  $K_c$  is a critical connectivity given by the curve

$$K_c 2p(1-p) = 1. \quad (6)$$

An equivalent derivation of (6) was also proposed more recently in Ref. [9]. One of the principal ingredients of both derivations is that the effect that the number of arguments, *i.e.* the connectivity  $K$ , excerpts over the Boolean functions  $\mathfrak{F}_i$  of (3) grows in average as orders  $2^K$ . This, while approximately true for  $K \gg 1$ , is rough when  $K \sim 1$  which are the typical values of  $K_c$  in (6) for the important zone  $p \sim 1/2$ ; as a recent study of the Boolean functions in terms of their irreducibility degree has shown [4].

In Ref. [4], a classification of Boolean functions according to the number of their arguments which really have an influence on the output of the functions; has been proposed. It was called the *irreducible degree* of the Boolean functions. By its means equations like (5) had been possible to calculate [3,4].

The scope of the present work is to make a mean field analysis of (3) taking into account the Boolean function irreducibility. As we will see, the transition Derrida-Stauffer curve (6) is going to appear, for each  $p$  at bigger values of  $K$ : particularly in the  $p \sim 1/2$  zone this effect will become more

pronounced, and will approached asymptotically to (6) in the extreme zones  $p \sim 0$  and  $p \sim 1$ .

The article is organized as follows: In Sec. 2 I write a combinatorial formulation for the dynamical equations (3), which allows to work in a frame more suitable for calculations. In Sec. 3 the concept of Boolean irreducibility is introduced, and quantitative expressions for the number of functions with a fixed degree of irreducibility are obtained. In Sec. 4 a mean field approach, taking into account Boolean irreducibility is established, and a new critical connectivity  $K_c$  is found. In Sec. 5 the average of the degree of irreducibility plus the probability of change of a Boolean function is calculated. This allows to obtain the corrected-for-irreducibility critical curve  $K_c$ . In Sec. 6 I set up my conclusions. In the appendixes, some properties of the combinatorial coefficients which are used in the work are quoted.

## 2. Combinatorial Notation for $NK$ -Automata

Let us write the evolution equation (3) in a formal language elaborated in Refs. [3,4]; which is more suitable to understand the combinatorial structure of the  $NK$ -Kauffman networks. Throughout the article:  $\forall S, S' \in \mathbb{Z}_2$ ,  $S+S' \in \mathbb{Z}_2$  is intended to be modulo 2. By  $[N] = \{1, 2, \dots, N\}$  I denote the set of the first  $N$  natural numbers.

### Definition 1:

Let

$$\mathfrak{C}_K^N = \left\{ C_K^{(\alpha)} \right\}_{\alpha=1, \dots, \binom{N}{K}}$$

denotes the collection of all the subsets of  $[N]$  with cardinality  $K$  ( $0 \leq K \leq N$ ), arranged in some unspecified order  $\alpha$ .

### Definition 2:

Each element  $C_K^{(\alpha)}$  of  $\mathfrak{C}_K^N$  is called a  $K$ -connection set, and so may be denoted by

$$C_K^{(\alpha)} = \{i_1, i_2, \dots, i_K\} \subseteq [N],$$

with,  $i_1 < i_2 < \dots < i_K$ ;  $i_m \in [N]$  ( $m = 1, \dots, K$ ).

### Definition 3:

To each  $K$ -connection set  $C_K^{(\alpha)}$  I associate a  $K$ -connection function

$$C_K^{*(\alpha)} : \mathbb{Z}_2^N \longrightarrow \mathbb{Z}_2^K \quad (7)$$

which projects each Boolean variable  $\mathbf{S} = (S_1, \dots, S_N) \in \mathbb{Z}_2^N$  onto  $\mathbb{Z}_2^K$ , and is defined by

$$C_K^{*(\alpha)}(\mathbf{S}) = C_K^{*(\alpha)}(S_1, \dots, S_N) = (S_{i_1}, \dots, S_{i_K}) \quad \forall \mathbf{S} \in \mathbb{Z}_2^N.$$

**Definition 4:**

A *K-Boolean function* is a map

$$b_K : \mathbb{Z}_2^K \rightarrow \mathbb{Z}_2, \quad (8)$$

and its negation  $\neg b_K$  is given by  $\neg b_K = b_K + 1$ .

**Definition 5:**

A *K-Boolean function* (8) is completely determined by its *truth table*  $\mathfrak{B}(b_K)$ , given by

$$\mathfrak{B}(b_K) = [\sigma_1, \sigma_2, \dots, \sigma_{2^K}], \quad (9)$$

where,  $\sigma_s \in \mathbb{Z}_2$ , is the  $s$ -th image of (8); and

$$s = s(\mathbf{S}) = 1 + \sum_{i=1}^K S_i 2^{i-1} \quad 1 \leq s \leq 2^K. \quad (10)$$

defines a **total order** among the possible  $2^K$  inputs of the argument  $\mathbf{S} \in \mathbb{Z}_2^K$  of the *K-Boolean function* (8).

There are  $2^{2^K}$  *K-truth tables*  $\mathfrak{B}(b_K)$ . Each *K-Boolean function* can be, uniquely classified according to Wolfram's notation by an integer number  $\mu = 1, \dots, 2^{2^K}$  given by <sup>[4,10]</sup>

$$\mu = 1 + \sum_{s=1}^{2^K} 2^{s-1} \sigma_s;$$

which also defines a **total order** among the *K-Boolean functions*.

I add a superscript  $\mu$  to each of the *K-Boolean functions* (8) and make the

**Definition 6:**

The set of *all K-Boolean functions* is given by

$$\Xi_K = \left\{ b_K^{(\mu)} : \mathbb{Z}_2^K \longrightarrow \mathbb{Z}_2 \right\}_{\mu=1}^{2^{2^K}}. \quad (11)$$

I clarify our abstract notation by the all important example of the truth table (9), for the case  $K = 2$ ; presented in Table 1. The first line represents their Wolfram's number  $\mu$ , indicated in boldface. At the bottom of the table:  $\mathfrak{F}$  stands for the logical meaning of each 2-Boolean function, with  $S_i$  ( $i = 1, 2$ ) representing the identity 2-Boolean function in the  $i$ -th argument of (10), while  $\neg S_i = S_i + 1$  its negation. The parameters  $\lambda$  and  $\omega$ , to be defined below, represent: the degree of irreducibility, of the corresponding 2-Boolean function, and its weight; respectively.

$\mathfrak{B}(b_2)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\sigma_1$	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
$\sigma_2$	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
$\sigma_3$	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
$\sigma_4$	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
$\mathfrak{F}$	$\neg\tau$	$\neg\vee$	$\nRightarrow$	$\neg S_2$	$\nLeftarrow$	$\neg S_1$	$\nLeftarrow$	$\neg\wedge$	$\wedge$	$\Leftrightarrow$	$S_1$	$\Leftarrow$	$S_2$	$\Rightarrow$	$\vee$	$\tau$
$\lambda$	0	2	2	1	2	1	2	2	2	2	1	2	1	2	2	0
$\omega$	0	1	1	2	1	2	2	3	1	2	2	3	2	3	3	4

**Table 1.** The  $\mathfrak{B}(b_2)$  truth tables of the sixteen 2-Boolean functions.

The maps (8) and (7) may be composed to represent the map (3) (for each Boolean variable  $S_i$ ) by

$$\mathbb{Z}_2^N \xrightarrow{C_K^{*(\alpha_i)}} \mathbb{Z}_2^K \xrightarrow{b_K^{(\mu_i)}} \mathbb{Z}_2 \quad i = 1, \dots, N;$$

with  $b_K^{(\mu_i)}$  and  $C_K^{*(\alpha_i)}$  extracted randomly according to the rules (1) and (2) respectively. So, the evolution equation (3) may be rewritten as

$$S_i(t+1) = b_K^{(\mu_i)} \circ C_K^{*(\alpha_i)}(\mathbf{S}(t)), \quad i = 1, \dots, N,$$

which defines an endomorphism  $\mathbb{Z}_2^N \rightarrow \mathbb{Z}_2^N$ ; but a particular one due to the presence of the  $K$ -connection map  $C_K^{*(\alpha_i)}$ . So, for  $K < N$ ;  $\mathcal{L}_K^N \subsetneq \mathcal{B}_N \cong \mathcal{G}_{2^N}$ , and only for the case  $K = N$  (of the random map model);  $\mathcal{L}_N^N \equiv \mathcal{B}_N \cong \mathcal{G}_{2^N}$  [3,4].

This is about rewriting (3) in a combinatorial notation. Now, for its random construction this is formally done by:

- i) According to (2): Extracting each  $C_K^{*(\alpha_i)}$  function (7), with equiprobability and without repetition, among the possible  $\binom{N}{K}$   $K$ -connection sets.

- ii) According to (1): Extracting each  $b_K^{(\mu_i)}$  from a probability distribution such that, in (9)  $\sigma_s = 1$ , ( $s = 1, \dots, 2^K$ ) with probability  $p$ ; which is attained by the probability distribution

$$\Pi_p(b_K) = p^\omega (1-p)^{2^K-\omega}, \quad (12)$$

where  $\omega = 0, 1, \dots, 2^K$  denotes the value of the *weight function*  $\omega(b_K)$  of  $b_K$ , defined by

$$\omega(b_K) = \sum_{s=1}^{2^K} \sigma_s. \quad (13)$$

The following disjoint decomposition is going to be important for the calculations of the next section:

$$\Xi_K = \bigsqcup_{\omega=0}^{2^K} \mathfrak{P}_K(\omega), \quad (14.a)$$

where

$$\mathfrak{P}_K(\omega) = \{b_K \in \Xi_K | \omega(b_K) = \omega\} \quad (14.b)$$

with cardinality

$$\#\mathfrak{P}_K(\omega) = \binom{2^K}{\omega}. \quad (14.c)$$

### 3. Irreducible Boolean Functions

In fact, not all of the  $K$ -Boolean functions depend strictly on their  $K$  arguments. For example, for  $K = 2$ , in Table 1: Rules **1** and **16** (*contradiction* and *tautology*, respectively) do not depend on neither  $S_1$  or  $S_2$ ; rules **4**, **6**, **11**, and **13** depend only on one of the arguments; while the remaining 10 depend on both  $S_1, S_2$ . Due to this fact, I do the following

#### Definitions 7

- i) A  $K$ -Boolean function  $b_K$  is *irreducible* on its  $m$ -th argument  $S_m$  ( $m = 1, \dots, K$ ), iff exist an  $\mathbf{S} \in \mathbb{Z}_2^K$  for which

$$b_K(S_1, \dots, S_m, \dots, S_K) = 1 + b_K(S_1, \dots, S_m + 1, \dots, S_K),$$

while, if this does not happen, the  $K$ -Boolean function  $b_K$  is *reducible* on the  $m$ -th argument  $S_m$ .

- ii) A  $K$ -Boolean function  $b_K$  is said to have a *degree of irreducibility*  $\lambda$  ( $\lambda = 0, 1, \dots, K$ ); if it is irreducible on  $\lambda$  of their arguments and reducible on the remaining  $K - \lambda$ .
- iii) If  $\lambda = K$ , the  $K$ -Boolean function is called to be *totally irreducible*.

Let  $\lambda(b_K)$  denotes the function that gives the *degree of irreducibility* of  $b_K$ , and  $\lambda$  ( $0 \leq \lambda \leq K$ ) their possible values. Then  $\Xi_K$  may, also be, disjoint decomposed by

$$\Xi_K = \bigsqcup_{\lambda=0}^K \mathfrak{T}_K(\lambda), \quad (15)$$

where

$$\mathfrak{T}_K(\lambda) = \{b_K \in \Xi_K \mid \lambda(b_K) = \lambda\}.$$

The cardinal coefficients  $\beta_K(\lambda) \equiv \#\mathfrak{T}_K(\lambda)$  where calculated recursively, in Ref. [4] obtaining the formula

$$\beta_K(\lambda) = \binom{K}{\lambda} \mathfrak{G}_\lambda, \quad (16)$$

where  $\mathfrak{G}_\lambda \equiv \beta_\lambda(\lambda)$  is the number of totally irreducible  $\lambda$ -Boolean functions. From (11), and taking cardinalities in (15); it can be shown that (16) obeys the recursion formula

$$2^{2^\lambda} = \sum_{\lambda=0}^K \binom{K}{\lambda} \mathfrak{G}_\lambda;$$

which may be inverted using Comtet's formulas <sup>[11]</sup> for the combinatorial coefficients (see Appendix A) obtaining

$$\mathfrak{G}_\lambda = \sum_{m=0}^{\lambda} (-1)^{\lambda-m} \binom{\lambda}{m} 2^{2^m}. \quad (17)$$

All coefficients  $\beta_K(\lambda)$ , but  $\beta_K(0) = 2$ , grow with  $K$ .  $\mathfrak{T}_K(0)$  consists exactly of the  $K$ -contradiction  $\neg\tau \equiv b_K^{(1)}$  and  $K$ -tautology  $\tau \equiv b_K^{(2^{2^K})}$  functions, with truth tables (9) given by

$$\mathfrak{B}(\neg\tau) = \underbrace{[0, 0, \dots, 0]}_{2^K}, \quad (18)$$

and

$$\mathfrak{B}(\tau) = \underbrace{[1, 1, \dots, 1]}_{2^K} \quad (19)$$



respectively.

On the other extreme, from (17) I obtain the asymptotic expression,

$$\frac{\mathfrak{G}_K}{2^{2^K}} \approx 1 - \mathcal{O}\left(\frac{K}{2^{2^{K-1}}}\right), \quad (20)$$

for  $K \gg 1$ . This shows that, with respect to the normalized counting measure, almost any  $K$ -Boolean function is irreducible.

These facts show us that the “real connectivity” of a  $K$ -Boolean function  $b_K$  is not  $K$ , but  $\lambda(b_K)$ . However for big values of  $K$  the “real connectivity” becomes nearly  $K$ . Even though the curve for the phase transition given by (6), in the region  $p \sim 1/2$  predicts values of the order  $K \sim 2$  so; in this case the effect of irreducibility should be appreciable in a mean field treatment that takes into account the degree of irreducibility. Let us calculate the average value  $\langle \lambda \rangle$  of  $\lambda$  as a function of  $p$  and  $K$ ; which represents the average connectivity of the automata (3) with respect to the extraction probability (12). So

$$\begin{aligned} \langle \lambda \rangle &= \sum_{b_K \in \Xi_K} \lambda(b_K) \Pi_p \circ \omega(b_K) = \sum_{\lambda=0}^K \lambda \sum_{b_K \in \mathfrak{T}_K(\lambda)} \Pi_p \circ \omega(b_K) \\ &= \sum_{\lambda=0}^K \lambda \sum_{\omega=0}^{2^K} \Pi_p(\omega) \sum_{b_K \in [\mathfrak{T}_K(\lambda) \cap \mathfrak{P}_K(\omega)]} 1 \\ &= \sum_{\lambda=0}^K \lambda \sum_{\omega=0}^{2^K} \Pi_p(\omega) \varrho_K(\lambda, \omega), \end{aligned} \quad (21)$$

where

$$\varrho_K(\lambda, \omega) = \# [\mathfrak{T}_K(\lambda) \cap \mathfrak{P}_K(\omega)].$$

The calculation of the cardinality  $\varrho_K(\lambda, \omega)$  is a difficult task that has been pursued in Ref. [12] using algebraic theoretical tools to do the combinatorial counting. Here I quote the result and refer the interested reader to the bibliography [12]:

$$\begin{aligned} \varrho_K(\lambda, \omega) &= \binom{K}{\lambda} \sum_{m=0}^K (-1)^{\lambda-m} \binom{\lambda}{m} \times \\ &\quad \times \delta([\omega 2^{m-K}] - \omega 2^{m-K}) \binom{2^m}{[\omega 2^{m-K}]}, \end{aligned}$$

where for all  $a \in \mathbb{R}$ ,

$$\delta(a) = \begin{cases} 1 & \text{if } a = 0 \\ 0 & \text{if } a \neq 0 \end{cases}$$

is Kronecker's delta,  $\lfloor a \rfloor \in \mathbb{Z}$  the *floor function*, which denotes the greatest integer  $\lfloor a \rfloor$  such that  $\lfloor a \rfloor \leq a$ ,  $\binom{\lambda}{m} = 0$ , for  $m > \lambda$ , and  $0^0 \equiv 1$ . With this aid, now it may be calculated (21), obtaining

$$\langle \lambda \rangle = K \left( 1 - [1 - 2p(1 - p)]^{2^{K-1}} \right) < K.$$

See Appendix **B** for manipulation of the combinatorial coefficients, and representations of Kronecker's delta in terms of them.

For  $p \sim 1/2$  the effect of (20) tends to dominate for  $K \gg 1$ , making  $\langle \lambda \rangle \sim K$ . Instead, for  $p$  near to 0 or 1, the  $K$ -Boolean function (18) or (19), respectively dominates (since contradiction and tautology are for each case the only functions to have a significant probability to be extracted); thus making  $\langle \lambda \rangle \sim 0$ . In Figure 1, a graph of  $\langle \lambda \rangle / K$  *vs.*  $p$  shows the behavior for different constant values of  $K$ .

#### 4. Mean Field Theory for $NK$ -Automata

Now a mean field approach ( $N \rightarrow \infty$ ) is developed to study the behavior of the Hamming distance of two initially nearby states  $\mathbf{S}$  and  $\mathbf{S}'$ , with respect to the parameters  $K$ , and  $p$  of the  $NK$ -Automata.

##### Definition 8:

The *Hamming distance*  $d_H$  between two states  $\mathbf{S}, \mathbf{S}' \in \mathbb{Z}_2^N$  is given by

$$d_H(\mathbf{S}, \mathbf{S}') = \sum_{i=1}^N (S_i + S'_i), \quad (22)$$

with the addition *modulo 2*. So, (22) counts the number of Boolean variables in  $S_i, S'_i \in \mathbb{Z}_2$  which differ, no matters their  $i$ -th position.

We want to see the evolution of  $d_H$  (22) as the system evolves in time according to (3) starting with two arbitrary, but nearby, states  $\mathbf{S}(0)$ ,  $\mathbf{S}'(0)$

at  $t = 0$ . Let us use the shorthand notation

$$d_H(t) \equiv \sum_{i=1}^N (S_i(t) + S'_i(t))$$

for the Hamming distance at time  $t$ . Let, also,  $d_H(0) \equiv \varepsilon N$  (with  $0 < \varepsilon \ll 1$ ) be the initial distance among them, with  $\varepsilon$  not scaling with  $N$ . The initial Hamming distance, is then such that

$$1 \ll d_H(0) \ll N,$$

so for  $N \gg 1$ , statistics may be done. Since the Boolean functions and their connections are randomly chosen from (1), and (2), respectively:

Each affected site  $i$ , such that  $S_i(0) \neq S'_i(0)$ ; will affect, on average,  $K$  sites  $\mathcal{A}_K \equiv \{i_1, \dots, i_K\} \subseteq [N]$ . The  $i_l$ -th affected site ( $l = 1, \dots, K$ ) is going to be the argument of a  $K$ -Boolean function  $b_K$  which is a stochastic variable obtained from the probability distribution (12). So,  $b_K$  is going to have a degree of irreducibility  $\lambda(b_K)$ , and a *probability of change*  $P_c(b_K)$  due that one of their arguments has changed (to be calculated later). Averaging over  $K$ , the  $i$ -th affected site will increase or decrease the Hamming distance at time  $t = 1$  by the stochastic factor

$$\Phi_i \equiv \frac{1}{K} \sum_{b_K \in \mathcal{A}_K} P_c(b_K) \lambda(b_K),$$

where  $0 \leq \Phi_i \leq K$ . At  $t = 1$  the Hamming distance grows (or decays), on average, according to

$$d_H(1) = \sum_{i=1}^N (S_i(1) + S'_i(1)) = \sum_{i=1}^N \Phi_i (S_i(0) + S'_i(0)),$$

Since all the sites  $i$  are equivalent in the definition (22) of the Hamming distance, I may order them so that the first  $\varepsilon N$  give the contribution to  $d_H(0)$  different from zero, obtaining

$$d_H(1) = \sum_{i=1}^{\varepsilon N} \Phi_i.$$

Due to the independence of each of the terms of the sum, I may apply the central limit theorem to, asymptotically, substitute  $\sum_i \Phi_i$ ; by its average, thus obtaining the asymptotic approximation

$$d_H(1) \approx \varepsilon N \langle \Phi \rangle \approx \varepsilon N \langle \lambda(b_K) P_c(b_K) \rangle, \quad (23)$$

where  $\langle \lambda(b_K) P_c(b_K) \rangle$  is going to be calculated in the next Section. The *relative error*  $\mathcal{E}_r$  in the calculation of (23), is given in terms of the standard deviation  $\sigma$ , of  $\langle \lambda(b_K) P_c(b_K) \rangle$ ; which, as long as  $\lambda(b_K) P_c(b_K)$  is a bounded stochastic variable; turns to be bounded. Then

$$\mathcal{E}_r = \frac{\sigma}{\langle \lambda(b_K) P_c(b_K) \rangle \sqrt{\varepsilon N}},$$

and so vanishes in the limit  $N \rightarrow \infty$ .

We may apply the same arguments for any  $t$  in the derivation of (23), as long as  $1 \ll d_H(t) \ll N$ , continues to be true, obtaining

$$d_H(t+1) \approx d_H(t) \langle \lambda(b_K) P_c(b_K) \rangle.$$

Solving for the initial condition  $d_H(0) = \varepsilon N$  we have, while  $1 \ll d_H(t) \ll N$  holds, the mean field equation for the evolution of Hamming distance

$$d_H(t) \approx \varepsilon N \exp \{t \ln [\Delta(K, p)]\}, \quad (24)$$

where

$$\Delta(K, p) = \langle \lambda(b_K) P_c(b_K) \rangle.$$

Now, from (24) we see that there is an exponential grow (or decay) in  $d_H(t)$  depending on whether  $\Delta(K, p)$  is bigger (or smaller) than one; which divides the phase space of the parameters  $p$  and  $K$  in the regions

$$\Delta(K, p) = \begin{cases} > 1 & \text{Standing for a disordered phase, called } \textit{chaotic}. \\ < 1 & \text{Representing an } \textit{ordered}, \text{ or } \textit{frozen phase}. \end{cases}$$

While

$$\Delta(K_c, p) = 1 \quad (25)$$

represents the equation for the phase space critical transition boundary.

## 5. Phase Space Diagram corrected for Boolean Decomposition

I now study for which values of the parameters  $K$ , and  $p$  (25) holds. This is done calculating the average  $\langle \lambda(b_K) P_c(b_K) \rangle$ . Note aboard that, due to the fact that  $0 \leq \langle \lambda(b_K) \rangle < K$ , and  $0 \leq \langle P_c(b_K) \rangle \leq 1$ :  $\Delta(K, p) < 1$  for

$K \leq 1$ . So  $K > 1$  is a necessary, but not sufficient, condition for chaotic behavior to be exhibited in  $NK$ -Kauffman networks.

The probability that a  $K$ -Boolean function  $b_K$  changes; due that one of its arguments has changed  $P_c(b_K)$  is, by definition, given by

$$P_c(b_K) = \sum_{\sigma \in \mathbb{Z}_2} \pi(b_K : \sigma) \pi(b_K : \sigma + 1 \mid \sigma), \quad (26)$$

where  $\pi(b_K : \sigma)$  is the probability to extract at random the value  $\sigma \in \mathbb{Z}_2$  from the images of the truth table (9)  $\mathfrak{B}(b_K)$ , and  $\pi(b_K : \sigma + 1 \mid \sigma)$  is the probability to extract at random the value  $\neg\sigma \equiv \sigma + 1$ , from  $\mathfrak{B}(b_K)$ ; given that  $\sigma$  has been previously extracted. Then, from (13)

$$\pi(b_K : \sigma) = \frac{\omega(b_K)}{2^K} \delta(\sigma + 1) + \frac{2^K - \omega(b_K)}{2^K} \delta(\sigma),$$

while

$$\pi(b_K : \sigma + 1 \mid \sigma) = \Sigma_K \pi(b_K : \sigma + 1),$$

where

$$\Sigma_K = \frac{2^K}{2^K - 1},$$

is a *second extraction factor* which appears since now there remain in the pool  $2^K - 1$  states  $\mathbf{S} \in \mathbb{Z}_2^N$  to choose. Substituting in (26) I obtain

$$P_c(b_K) \equiv P_c \circ \omega(b_K) = 2 \Sigma_K \frac{\omega(b_K) (2^K - \omega(b_K))}{2^{2K}},$$

Now I may calculate  $\langle \lambda P_c \rangle \equiv \langle \lambda(b_K) P_c(b_K) \rangle$  much in the same way as (21):

$$\begin{aligned} \langle \lambda P_c \rangle &= \sum_{b_K \in \Xi_K} \Pi_p \circ \omega(b_K) \lambda(b_K) P_c \circ \omega(b_K) \\ &= \sum_{\omega=0}^{2^K} \Pi_p(\omega) P_c(\omega) \sum_{\lambda=0}^K \lambda \varrho_K(\lambda, \omega). \end{aligned}$$

Now, with by the aid of (B3) and (B5) of Appendix **B**

$$\langle \lambda P_c \rangle = 2 K p (1 - p) \left\{ 1 - 2 p (1 - p) [1 - 2 p (1 - p)]^{2^{K-1}-2} \right\}.$$

is obtained.

So the critical transition curve (25) is given by

$$2 K_c p (1 - p) \left\{ 1 - 2 p (1 - p) [1 - 2 p (1 - p)]^{2^{K_c - 1} - 2} \right\} = 1. \quad (27)$$

Comparison with Derrida-Stauffer's result (6) shows the appearance of a new factor of order  $\mathcal{O}[1 - 2 p (1 - p)]$  which accounts for the existence of irreducibility in the  $K$ -Boolean functions. Figure 2 compares the graphs of (6) and (27). Now, due to the effect of irreducibility the transition occurs for each  $p$  at greater values of  $K$ . In particular, for the all important case  $p = 1/2$ ;

$$K_c = 2.62140224613 \dots$$

## 6. Conclusion

I have re-calculated the phase transition Derrida-Stauffer's curve for the  $NK$ -Kauffman networks <sup>[8]</sup> by correcting the calculations for the effect of Boolean irreducibility in  $K$ -Boolean functions <sup>[4]</sup>. While it turns out not to be a big correction for this case, it is an important effect in many other aspects in the behavior of  $NK$ -Kauffman networks, such as the injective properties of the function  $\Psi$  (4), which maps the  $NK$ -Kauffman networks set  $\mathcal{L}_K^N$  into the  $N$ -Boolean functions set  $\mathcal{B}_N$  <sup>[3,4]</sup>.

Without doubt the degree  $\lambda$  of irreducibility, *Definition 7 (ii)* should play an important role in the characterization of the  $NK$ -Kauffman networks dynamics as a function of the parameters  $N$ ,  $K$ , and  $p$ . Work is in progress with these new tool.

## Acknowledgments

This work is supported in part by **PAPIIT** projects Nos. **IN101309-3** and **IN102712-3**. The author wish to thank: Alberto Verjovsky for fruitful mathematical discussions. Thalía Figueras for careful reading of the manuscript, Pilar López Rico for accurate services on informatics, Mariana Zertuche for data reduction, and Diego Hernández for computer's advise.

## Appendix A: Inversion Formula for Binomial Coefficients

In Comtet's work, the following inversion formula is proved <sup>[11]</sup>:

For any two sequences of real numbers

$$\{f_r\}_{r=0}^n, \text{ and } \{g_r\}_{r=0}^n, \quad n \geq 0$$

such that

$$f_n = \sum_{r=0}^n \binom{n}{r} g_r.$$

Then, it follows that the  $g_r$  are given in terms of the  $f_r$  through

$$g_n = \sum_{r=0}^n (-1)^{n-r} \binom{n}{r} f_r.$$

## Appendix B: Identities and checks for manipulating $\varrho_K(\lambda, \omega)$

It is useful for the calculations involving  $\varrho_K(\lambda, \omega)$  to extend the definition of the combinatorial coefficients when the upper index  $a \in \mathbb{R}$ , and the lower index  $n \in \mathbb{Z}$ ; by writing <sup>[13]</sup>:

$$\binom{a}{n} = \begin{cases} \frac{a(a-1)\dots(a-n+1)}{n!} & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases}. \quad (B1)$$

Which for the case  $a \in \mathbb{Z}$  gives  $\binom{a}{n} = 0$  if  $a < n$ . From (B1), the following identity can be proved to hold, for any  $a \in \mathbb{R}$ ,  $m, n \in \mathbb{Z}$  <sup>[13]</sup>

$$\binom{a}{m} \binom{m}{n} = \binom{a}{n} \binom{a-n}{m-n}. \quad (B2)$$

The *Binomial Theorem* for  $z \in \mathbb{C}$  comes to be

$$(1+z)^a = \sum_{m \geq 0} \binom{a}{m} z^m \quad \text{if} \quad \begin{cases} |z| < 1, & \text{for } a \in \mathbb{R} \\ \forall z \in \mathbb{C} & \text{for } a \in \mathbb{N} \cup \{0\} \end{cases} \quad (B3)$$

where,  $0^0 \equiv 1$ . From (B3) the following useful Kronecker's delta representations, for  $a \in \mathbb{R}$ , and  $a \geq 0$  may be obtained by deriving with respect to  $z$  and taking  $z = -1$  [where it is to be noted that the series (B3) still

converges for  $z = -1$  due to the alternating sign, and that  $\binom{a}{m} \sim \mathcal{O}(m^{-1-a})$ , for  $m \gg 1$  [13];

$$\delta(a) = \sum_{m \geq 0} (-1)^m \binom{a}{m}, \quad \delta(a-1) = \sum_{m \geq 0} (-1)^{m+1} m \binom{a}{m} \quad a \geq 0. \quad (B4)$$

Using (B2) and (B4) one easily obtain the following check identities for  $\varrho_K(\lambda, \omega)$ , which are consequences of (14), (16) and (17):

$$\sum_{\lambda=0}^K \varrho_K(\lambda, \omega) = \binom{2^K}{\omega},$$

and

$$\sum_{\omega=0}^{2^K} \varrho_K(\lambda, \omega) = \binom{K}{\lambda} \sum_{m=0}^{\lambda} (-1)^{m-\lambda} \binom{\lambda}{m} 2^{2^m} \equiv \beta_K(\lambda).$$

From (B3) and (B4) also follows the useful identity

$$\sum_{\lambda=0}^K \lambda \varrho_K(\lambda, \omega) = K \left[ \binom{2^K}{\omega} - \binom{2^{K-1}}{\lfloor \frac{\omega}{2} \rfloor} \delta\left(\lfloor \frac{\omega}{2} \rfloor - \frac{\omega}{2}\right) \right] \quad (B5)$$

## References

- [1] S.A. Kauffman, *Metabolic Stability and Epigenesis in Randomly Connected Nets*. J. Theoret. Biol. **22** (1969) 437.
- [2] S.A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press (1993).
- [3] D. Romero, and F. Zertuche, *Number of Different Binary Functions Generated by NK-Kauffman Networks and the Emergence of Genetic Robustness*. J. Math. Phys. **48** (2007) 083506.
- [4] F. Zertuche, *On the robustness of NK-Kauffman networks against changes in their connections and Boolean functions*. J. Math. Phys. **50** (2009) 043513.



- [5] D. Romero, and F. Zertuche, *The Asymptotic Number of Attractors in the Random Map Model*. J. Phys. A: Math. Gen. **36** (2003) 3691; *Grasping the Connectivity of Random Functional Graphs*. Stud. Sci. Math. Hung. **42** (2005) 1.
- [6] B. Derrida, and H. Flyvbjerg, *The Random Map Model: a Disordered Model with Deterministic Dynamics*. J. Physique **48** (1987) 971.
- [7] H. Flyvbjerg, and N.J. Kjaer, *Exact Solution of Kauffman's Model with Connectivity One*. J. Phys. A: Math. Gen. **21** (1988) 1695.
- [8] B. Derrida, and D. Stauffer, *Phase Transitions in Two-Dimensional Cellular Automata*. Europhys. Lett. **2** (1986) 739.
- [9] M. Aldana, S. Coppersmith, and L. Kadanoff, *Boolean Dynamics with Random Couplings*. In: Perspectives and Problems in Nonlinear Science, 23–89. Springer Verlag, New York (2003).
- [10] G. Weisbuch, *Complex Systems Dynamics*. Addison Wesley, Redwood City, CA (1991); Wolfram, S., *Universality and Complexity in Cellular Automata*. Physica D **10** (1984) 1.
- [11] L. Comtet, *Advanced Combinatorics*. Reidel, (1974), p. 165.
- [12] M. Takane and F. Zertuche.  $\mathbb{Z}_2$ -Algebras in the Boolean Function Irreducible Decomposition. Journal of Mathematical Physics **53**, 023516 (2012).
- [13] R.L. Graham, D.E. Knuth, and O. Patashnik, *Concrete Mathematics*, New York: Addison-Wesley (1994).

### Figure caption

Figure 1. (Color online), Graph of  $\langle \lambda \rangle / K$  vs.  $p$ , for constant values of  $K$ . It shows the behavior of the *average degree of irreducibility* of the  $K$ -Boolean functions  $\Xi_K$  with respect to the bias  $p$ .

Figure 2. (Color online) Compares the graphs of the Derrida-Stauffer phase transition curve (6), which does not take into account the irreducible degree of  $K$ -Boolean functions, with the graph of (27) which takes into account this effect.